HPC Matters!

How Supercomputing Impacts NASA's Mission

Dr. Piyush Mehrotra

Chief, NASA Advanced Supercomputing (NAS) Division

NASA Ames Research Center in Silicon Valley

http://www.nas.nasa.gov/hecc

HPC User Forum

June 22, 2022

2022 Strategic Plan Overview

NASA's Strategic Plan is organized around 4 themes and their related Strategic Goals.

Vision

Exploring the secrets of the universe for the benefit of all.

Mission

NASA explores the unknown in air and space, innovates for the benefit of humanity, and inspires the world through discovery.

Strategic Goals

- **DISCOVER** — EXPAND HUMAN KNOWLEDGE THROUGH NEW SCIENTIFIC DISCOVERIES.
- **EXPLORE** — EXTEND HUMAN PRESENCE TO THE MOON AND ON TOWARDS MARS FOR SUSTAINABLE LONG-TERM EXPLORATION, DEVELOPMENT, AND UTILIZATION.
- **DEVELOP** — CATALYZE ECONOMIC GROWTH AND DRIVE INNOVATION TO ADDRESS NATIONAL CHALLENGES.
- **ENABLE** — ENHANCE CAPABILITIES AND OPERATIONS TO CATALYZE CURRENT AND FUTURE MISSION SUCCESS.

Supercomputing Facility @ NASA Ames

NASA's Premier Supercomputer Center

Resources have broad mission impact across all of NASA's Missions

Over 600 science & engineering projects with more than 1,600 users

AITKEN

Vital Stats

3,200-node HPE E-Cell/Apollo 9000 system

308,224 cores total

13.1 petaflops theoretical peak

6.39 petaflops sustained performance (Nov. 2021)

1.27 petabytes total memory

ELECTRA

Vital Stats

3,456-node HPE ICE X/HPE E-Cell system

124,416 cores total

8.32 petaflops theoretical peak

5.44 petaflops sustained performance (June 2021)

589 terabytes total memory

PLEIADES

Vital Stats

11,207-node HPE ICE supercluster

241,324 cores total

7.09 petaflops theoretical peak

5.95 petaflops sustained performance (June 2021)

927 terabytes total memory

VISUALIZATION

Vital Stats

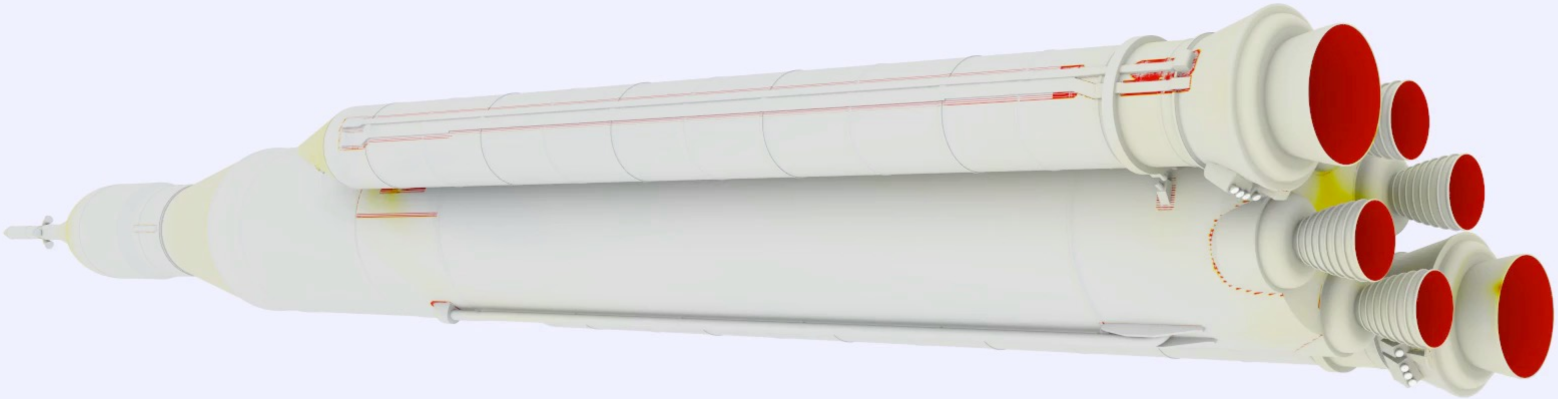
128-screen tiled LCD wall arranged in 8x16 configuration (23-ft. wide by 10-ft. high)

2,560 Intel Xeon Ivy Bridge processor cores

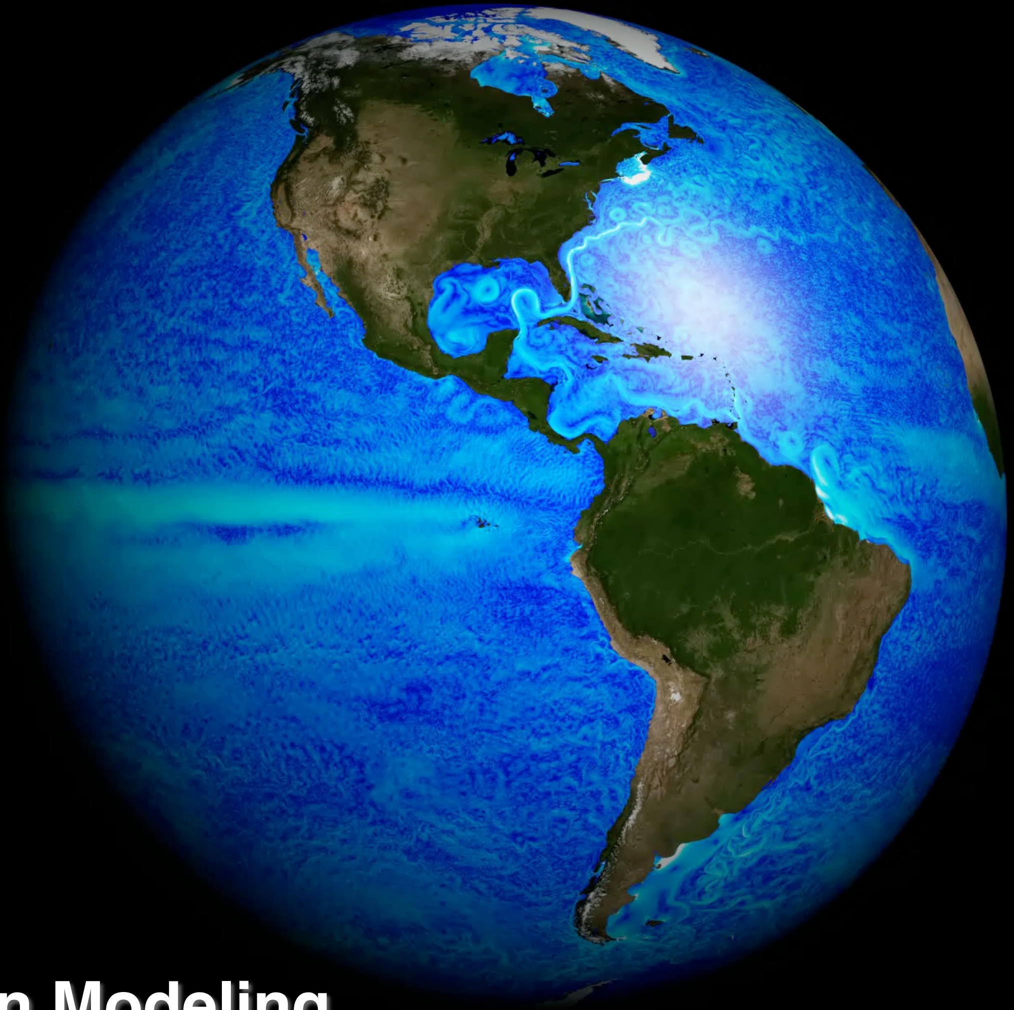
128 Nvidia GeForce GTX 780 Ti graphics processing units

Modeling the Launch Environment

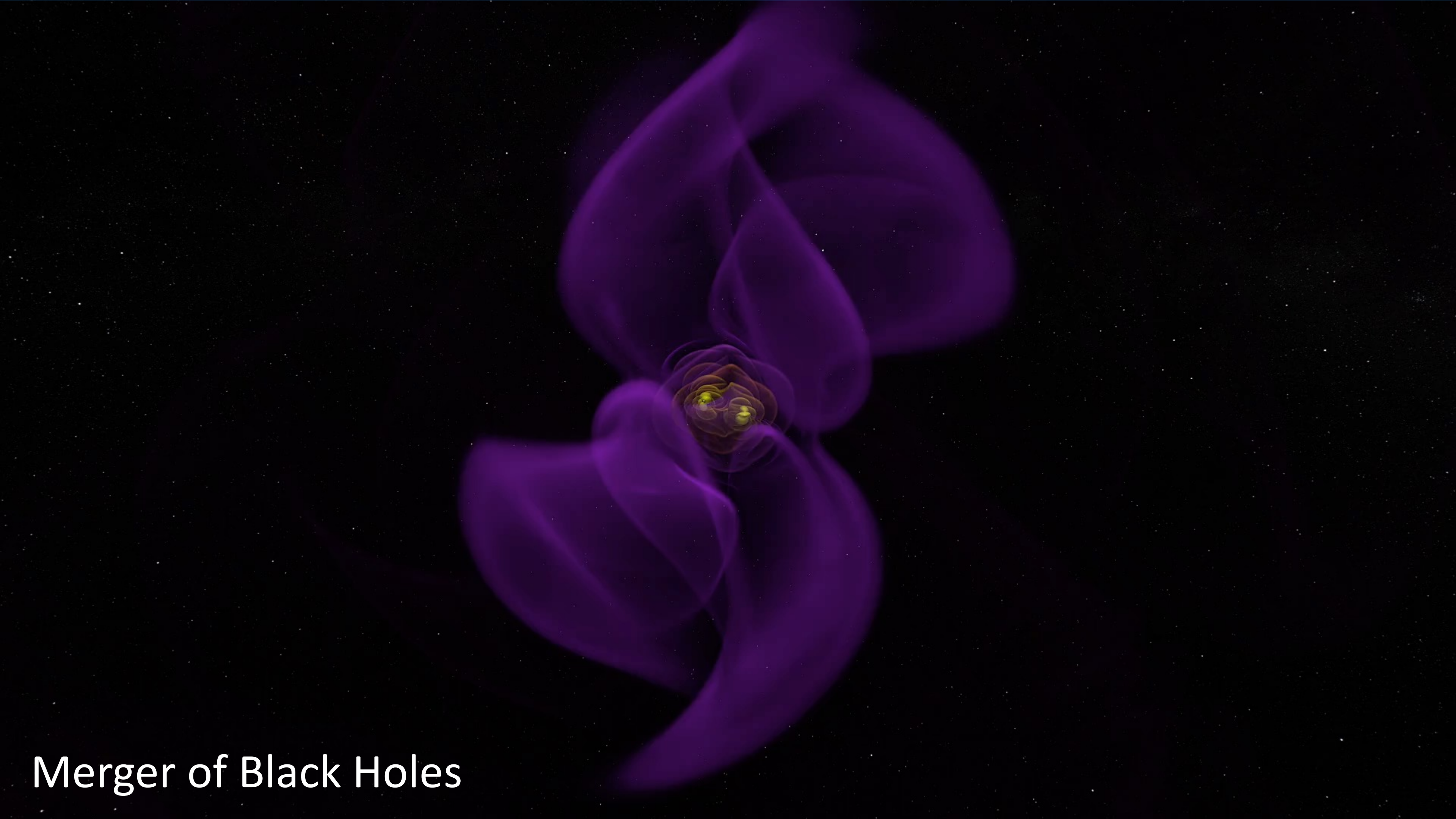
Space Launch System – Stage Separation



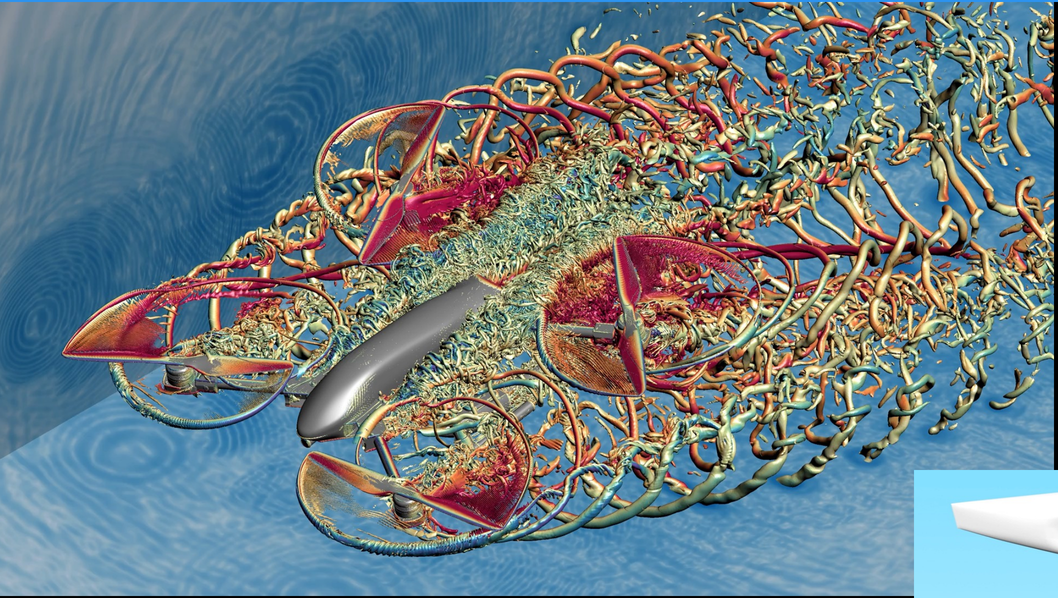
0 ' 0 "

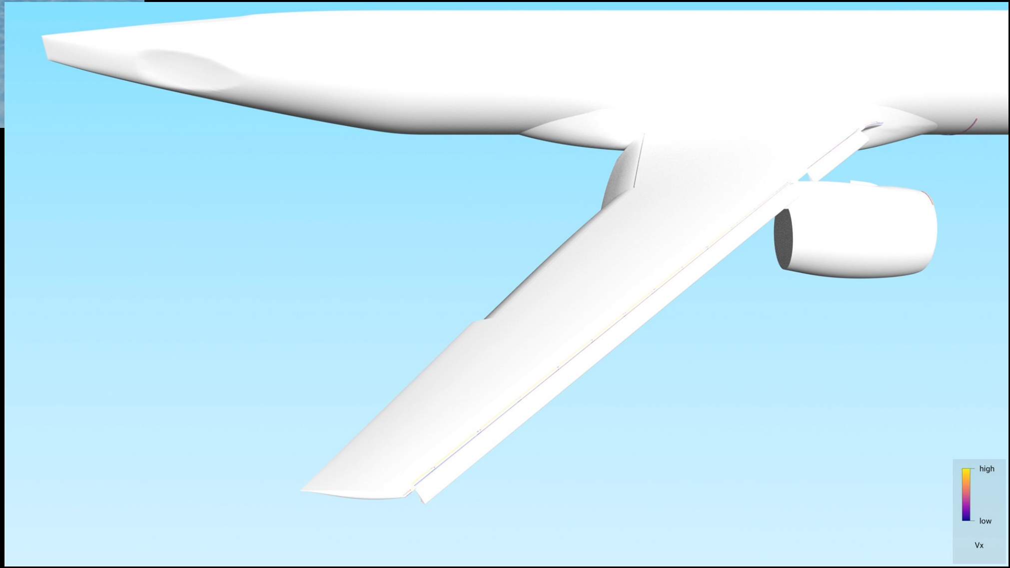


Global Ocean Modeling



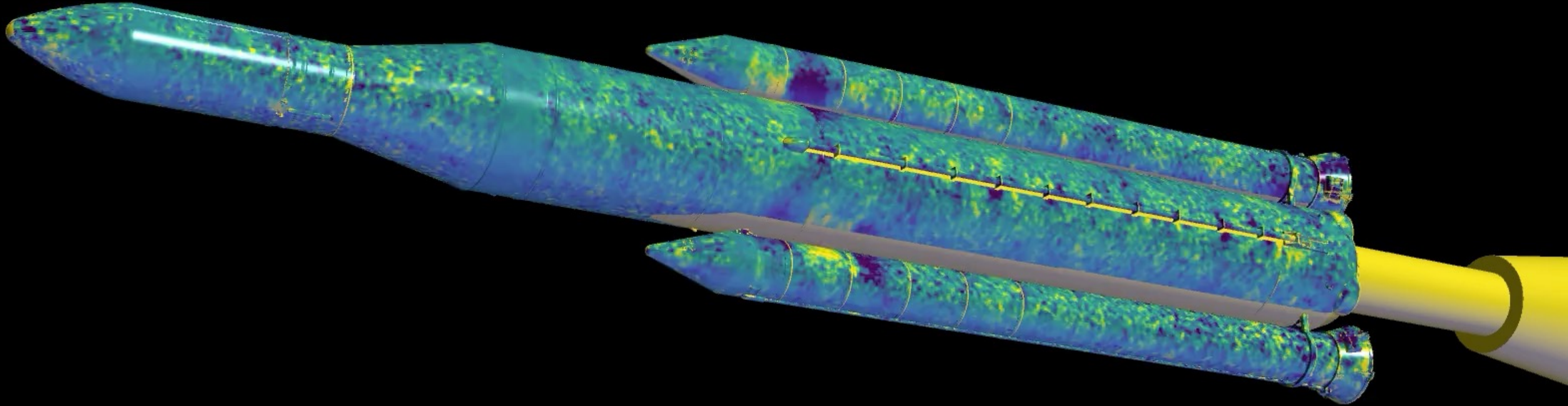
Merger of Black Holes





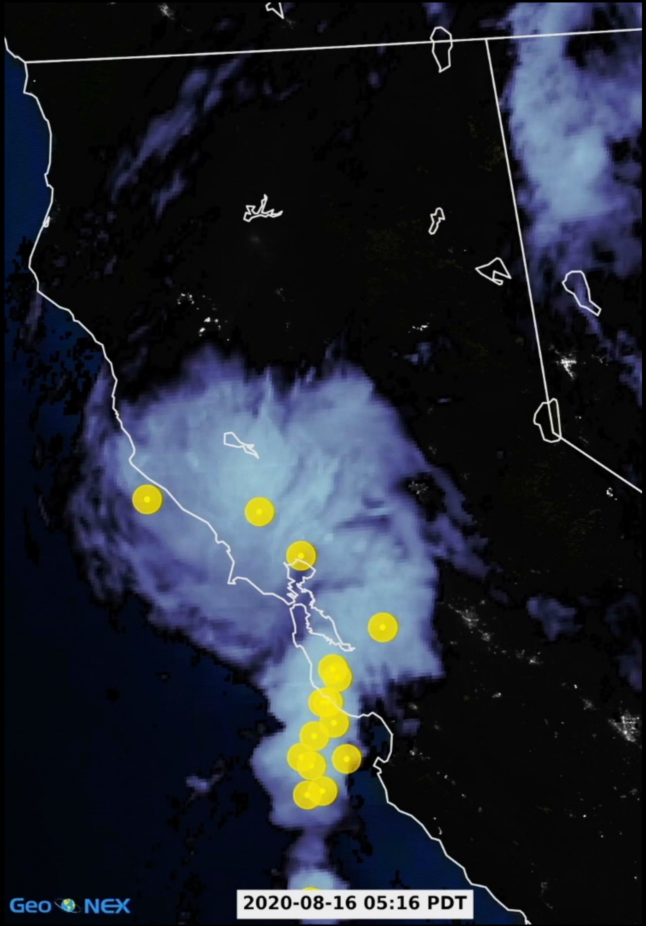
Aerosciences

Wind tunnel runs of SLS model using pressure sensitive paint to estimate pressure/loads



Near real-time analysis of test data using Supercomputer resources
with potential for computer-guided data acquisition

Near real-time analysis of from GOES-ABI Satellite Data for fire detection/evolution



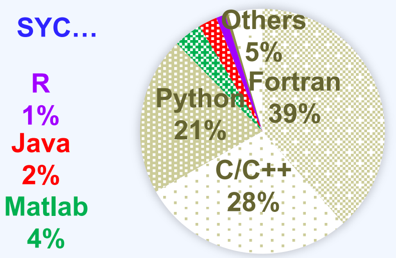
Sample NASA AI/ML Projects

- **Feature detection**
  - Shock waves & vortices from flow data
  - Exoplanet identification from TESS/Kepler data
  - Artifact identification from satellite data, e.g., trees, irrigated lands
  - Shock waves & vortices from flow data
- **Prediction**
  - Solar flares/space weather from solar surface magnetic fields data
  - Asteroid properties from light curves
  - Solar cells current-voltage properties from IV curves
- **Anomaly detection**
  - Aviation safety issues from flight data
  - Systems behavior, e.g., ISS control operations
- **Interactive ISS crew assistants/robots that can learn**
- **Autonomous rovers**
- **Machine Learning Emulators of Physics-based models**

- **Mission Support**
  - Email Classification/Records Management
  - Scientific Document Tagging
  - Network Traffic Anomaly Detection
  - Service Desk Ticket Analysis & Trending
  - Detect CUI content in documents

Programming Languages, Libraries, Commercial Software (2020 User Survey)

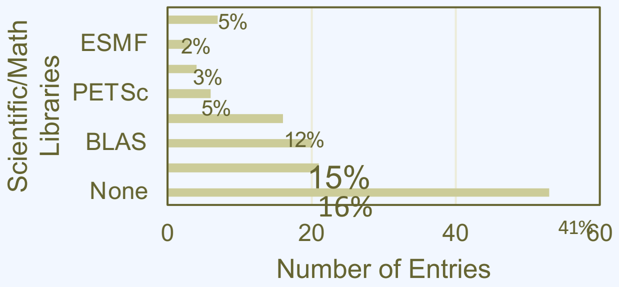
Programming Languages
(244 entries)



Others:
- Ruby (3 entries)
- Julia (2)
- CUDA/OpenMP (1)
- IDL (1)
- Tcl/tk (1)
- Shell scripting (1)
- Don't know (2)

➢ Fortran/C/C++ still dominate.
➢ Python is getting popular.
➢ SYCL/DPC++ is being explored (by FUN3D developers).

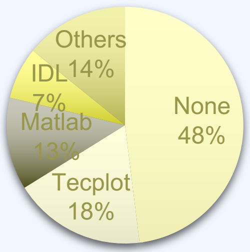
Scientific/Math Libraries
(130 entries)



Others:
- Armadillo (1)
- HYPRE, SLUG (1)
- Intel C runtime (1)
- Python (1)
- Don't know (3)

➢ 59% of entries use sci/math libraries.
➢ Intel MKL, BLAS, FFTW dominate.

Commercial Software
(127 entries)



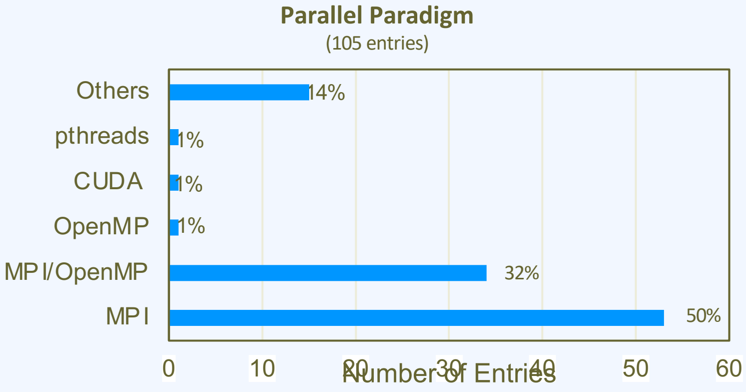
Other commercial: (8)
- Paraview (2)
- Powerflow (2)
- ANSA (1)
- CAMRADII (1)
- Pointwise (1)
- Totalview (1)

Non-commercial listed: (6)
- FITS, git, miniconda, netcdf,
- Python (2), tensorflow

Don't know: (3)

➢ Licensed Tecplot/Matlab/IDL still in need.
➢ Open source software packages are popular.

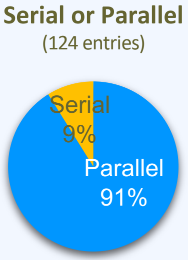
Parallelism in Applications (2020 User Survey)

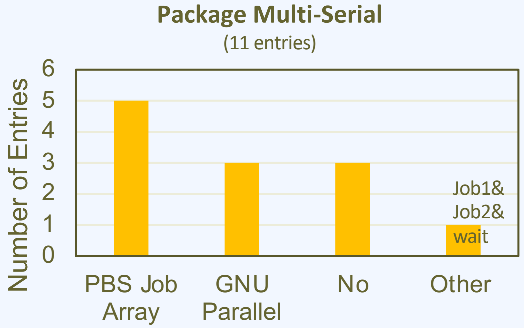


Others:

- Combination of
  - MPI/CUDA
  - MPI/OpenMP
  - MPI/OpenACC
  - MPI/pthreads
- SYCL
- Linda
- GNU Parallel (w/o MPI)
- OpenMP/Python multiprocessing

- MPI still dominates (~ 82% MPI or MPI/OpenMP).
- Pure OpenMP or pthreads not heavily used.
- CUDA programming begins to show up at HECC.
- Some interests in different hybrid parallelism: especially, MPI or MPI/OpenMP on CPU and MPI/CUDA on GPU.





- Most applications (91%) are parallel.
- For serial applications, packaging multi-serial is mostly done with Job Array or GNU Parallel.

Programming Challenges

- **Complex target hardware architectures/environments**
  - CPUs with increasing number of cores, deep memory hierarchies; accelerators; vector engines, GPUs, FPGAs, heterogeneous environments, complex I/O infrastructure
- **Multitude of programming models and environments**
  - Programming languages and libraries: C/C++, Fortran, OpenMP, MPI
  - Multiple levels of parallelism
  - Offload for accelerators: OpenACC, OpenMP target, NVIDIA CUDA, AMD HIP, Intel oneAPI, SYCL
  - Scripting languages and frameworks: Python, Julia, R, Kokkos, Raja
  - Domain-specific application frameworks and libraries
- **Users want both code and performance portability**
- **Large legacy code-bases**
  - Optimize existing code with some restructuring of code and data structures
  - Major rewrite to match architectures
  - Use different/more appropriate algorithms
- **Lack of budget and expert labor resources**

Conclusions

- Today's supercomputers are enabling ever larger simulations – using tens of thousands to hundreds of cores running for days even weeks.
- The increased computing capability has allowed for a dramatic increase in the fidelity of the simulations and the ensuing results
- The enhanced quality and granularity of the data has supported the decision makers, increased our understanding of the Earth system, solar system and the universe while also having a direct impact on our daily lives.

Supercomputing plays a key role
in support of
all of NASA's goals and objectives.

Questions?

piyush.mehrotra@nasa.gov

http://www.nas.nasa.gov/hecc